Hindawi
Complexity
Volume 2018, Article ID 9496731, 17 pages
https://doi.org/10.1155/2018/9496731

Research Article

Exploring Project Complexity through Project Failure Factors: Analysis of Cluster Patterns Using Self-Organizing Maps

Vicente Rodríguez Montequín, Joaquín Villanueva Balsera, Sonia María Cousillas Fernández, and Francisco Ortega Fernández

Project Engineering Area, University of Oviedo, C/Independencia 13, 33004 Oviedo, Spain

Correspondence should be addressed to Vicente Rodríguez Montequín; montequi@uniovi.es

Received 27 October 2017; Accepted 2 April 2018; Published 22 May 2018

Academic Editor: Luis Carral Couce

Copyright © 2018 Vicente Rodríguez Montequín et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the field of project management, complexity is closely related to project outcomes and hence project success and failure factors. Subjectivity is inherent to these concepts, which are also influenced by sectorial, cultural, and geographical differences. While theoretical frameworks to identify organizational complexity factors do exist, a thorough and multidimensional account of organizational complexity must take into account the behavior and interrelatedness of these factors. Our study is focused on analyzing the combinations of failure factors by means of self-organizing maps (SOM) and clustering techniques, thus getting different patterns about the project managers perception on influencing project failure causes and hence project complexity. The analysis is based on a survey conducted among project manager practitioners from all over the world to gather information on the degree of influence of different factors on the projects failure causes. The study is cross-sectorial. Behavioral patterns were found, concluding that in the sampled population there are five clearly differentiated groups (clusters) and at least three clear patterns of answers. The prevalent order of influence is project factors, organization related factors, project manager and team members factors, and external factors.

1. Introduction

As projects have become more and more complex, there is an increasing concern about the concept of project complexity and its influence upon the project management process. Projects have certain critical characteristics that determine the appropriate actions to manage them successfully. Project complexity (organizational, technological, informational, etc.) is one such project dimension. The project dimension of complexity is widespread within project management literature. In the field of projects, complexity is closed related to the causal factors to get the project outcomes, which in the project management field is usually referred to as "project success/failure factors". Better understanding of project success/failure factors is a key point for creating a strategy to manage complexity.

Most attributes of complexity are known to be constantly changing variables such as project type, project size, project location, project team experience, interfaces within a project, logistics/market conditions, geopolitical and social issues, and permitting and approvals [1]. Several studies focus on project complexity and the factors that influence its effect on project success. In general, there is not a clear difference between complexity and success factors when considering the literature of project management. For instance, Gidado defined project complexity and identified the factors that influence its effect on project success in relation to estimated production time and cost, based on literature search and structured interviewing of practitioners [2]. Kermanshachi et al. [3] consider that when complexity is poorly understood and managed, project failure becomes the norm. They focused on strategies to manage complexity in order to increase the likelihood of project success. Vidal and Marle [4] define project complexity as a property of a project, which makes it difficult to understand, foresee, and control the project's overall behavior. Remington et al. [5] believe that a complex project demonstrates a number of characteristics to a degree, or level of severity, that makes it difficult to predict

project outcomes or manage projects. One of the project complexity definitions that best fits the aim of this study was the one given by Kermanshachi et al.: "Project complexity is the degree of interrelatedness between project attributes and interfaces, and their consequential impact on predictability and functionality" [3]. This definition can guide the identification of project complexity indicators and management strategies which reduce the undesired outcomes often related to project complexity. Under this context, we can expect a relationship among project failure causes (derived from project attributes and interfaces) and project complexity.

Success in projects is more complex than just meeting cost, deadlines, and specifications. In fact, customer satisfaction with the final result has a lot to do with the perception of success or failure in a project. In the end, what really matters is whether the parties *associated with* and *affected by* a project are satisfied or not [6]. Meeting deadlines and costs does not really matter if the final project outcomes do not meet expectations.

However, our work is not focused on the concepts of success or failure but on the study of the aspects that lead to the failure in projects and their combinations, considering that the complexity of a project is determined by these factors. There are plenty of factors whose application is significant for the success or failure of a project. In the literature, these are called *critical success factors*, and many studies have been devoted to define, clarify, and analyze such factors. Success factors are subjected to the perceptions of the ones involved in the project development, depending not only on the stakeholder but also on cultural or geographical differences, which are reflected in the context of the organization [7]. There are also a lot of sectorial influences. For example, Huysegoms et al. have identified the causes of failure in software product lines context [8]. Other examples can be found in [9–12].

Obviously, projects fail due to many different reasons, if we understand "failure" as the systematic and widespread noncompliance of the criteria which define a successful project [13]. Nevertheless, due to the inner subjectivity of the concept, each person working in the same project has a personal opinion about the determining causes of its failure. These opinions can also vary depending on the type and sector of projects, so that distinctive patterns of causes are associated with the failure of specific kinds of projects. The most usual is that a combination of several factors with different levels of influence in different stages of the life cycle of a project result in its success or failure.

Most of the studies in the literature focus on determining lists of factors and their categorization, ranking them according to their influence level. Nevertheless, projects behavior derives from "systematic interrelated sets of factors" rather than single causal factors and the fact that true causes of project outcomes are difficult to identify [14, 15]. The interactions between the different factors seem to be as important as each factor separately. However, there seems to be no formal way to account for these interrelations, which may be the reason why this point is weakly treated in the literature. Our study is focused on analyzing the combination of these factors by means of clustering techniques, thus getting different patterns about the project manager practitioners perception on influencing project failure causes and hence project complexity.

The work here presented is part of a more global study analyzing success factors and failure causes in projects. A questionnaire has been aimed specifically to project management practitioners to gather information on the degree of influence of different factors on the failure or success in a project. The questionnaire inquires about their perception on the most influential factors to be considered to reach success, as well as the common failure causes they have most frequently encountered. A selection of critical success factors and failure causes were selected as a basis for the questionnaire, compiling previous research work results [16, 17] with the most frequent causes reflected in the literature. The questionnaire is generic, not intended for any specific sector or geographical area. Although it is not focused on any particular project or field, it gathers this type of information to be able to correlate it. The survey was distributed anonymously to recipients through LinkedIn, an Internet professional network. Such study determines the most frequent failure causes and the most important success factors in real world projects. In the initial stage, a statistical analysis of the sample data was conducted with the aim of answering the question of whether the valuations depend on the geographical areas of the respondents or on the types of projects that have been carried out [16]. The study shows that there is no absolute criterion and that subjectivity is the inherent characteristic of those valuations. So, as a complementary study, clustering techniques are applied in order to find patterns in the set of received answers. This paper presents the obtained results. Our study is a cross-sectorial study, which is a remarkable improvement over the majority of studies that consider only one sector, with the construction one being of the most studied.

The remainder of this paper is organized as follows. Section 2 reviews the literature related to the environment of this work. Section 3 describes the research methodology. Sections 4 and 5 present and discuss the results and finally Section 6 exposes the conclusions.

2. Literature Review

The literature is explored according to the three main points of this work: background of complexity in the field of project management, project success/failure causes and their connection with project complexity, and applications of self-organizing maps (SOM) in the field of project management.

2.1. Project Complexity Theory. Complexity has been recognized as one of the most relevant topics in project management research. According to Baccarini [18], one of the first reviews about complexity in the project management field, project complexity can be defined in terms of differentiation and interdependency and it is managed by integration. In the definition, differentiation refers to the number of varied components of the project (tasks, specialists, subsystems, and parts), and interdependency refers to the degree of interlinkages among these components. He established the dichotomy considering that project complexity is composed

of technological complexity and organizational complexity. Complexity can take various forms, namely social, technological, environmental, and organizational. Worth mentioning here is the work of Bosch-Rekveldt et al. [19] who proposed the TOE framework, consisting of fifty factors in three families: technical, organizational, and environmental. The authors have also concluded that organizational complexity worried project managers more than technical or environmental complexities. Vidal and Marle [4] argued that approximately 70% of project complexity factors are organizational. We follow these assertions by focusing our study mainly on the organizational factors, knowing that this covers a major area of complexity in projects.

Scholars have focused on the identification of complexity attributes more than any other topic in the field of project complexity. Studies in this area have evolved significantly over the past twenty years. Cicmil et al. [20] identified complexity as a factor that helps determine planning and control practices and hinders the identification of goals and objectives, or a factor that influences time, cost, and quality of a project. Gidado [2] defined project complexity and identified the factors that influence its effect on project success. Also, the study proposes an approach that measures the complexity of the production process in construction.

Project complexity has been analyzed in the book edited by Cooke-Davies [21] from three different perspectives: people who manage programs and projects (practitioners), line managers in organizations to which programs and projects make a substantial contribution (managers), and members of the academic research community who have an interest in how complexity shapes and influences the practice of program and project management (researchers). The book constitutes a valuable resource to put together what is currently known and understood about the topic, to help practitioners and their managers improve future practice, and to guide research into answering those questions that will best help to improve understanding of the topic.

Although most authors emphasized the influence of interdependencies and interactions of various elements on project complexity [22], few works analyze those dependencies. For example, TOE model [19] does not allow for an understanding of how various elements contribute to overall complexity. Nevertheless, other authors regarded project complexity as having nonlinear, highly dynamic, and emerging features. Vidal et al. [23], for example, proposed the definition of project complexity as "the property of a project which makes it difficult to understand, foresee, and keep under control its overall behavior, even when given reasonably complete information about the project system". Lu et al. [24] propose that project complexity can be defined as "consisting of many varied interrelated parts, and has dynamic and emerging features". Lessard et al. built the House of Project Complexity (HoPC), a combined structural and process-based theoretical framework for understanding contributors to complexity.

The connection between failure and complexity in project management has also been established in the literature. Ivory and Alderman [22] studied the failure in complex systems in order to shed some critical light on the management of complex projects. The authors conclude that it was not the technical complexity per se that made the management of the projects complex. Rather, the primary determinants of the complexity of the project management process stemmed from changes in the markets, regulatory context, and knowledge requirements facing the project.

There are connections between the project complexity indicators identified in the literature and the project failure causes. For example, Kermanshachi et al. [3] identified 37 indicators. Among them, several examples usually found in the literature as project failure causes are included (i.e., impact of external agencies on the project execution plan, impact of required approvals from external stakeholders, level of difficulty in obtaining permits, level of project design changes derived by Request for Information (RFI), etc.). In fact, there are several examples of authors that are referenced frequently in the field of both project complexity and project success factors. Shenhar et al. [25], Dvir and Lechler [26], Cooke-Davies [27], and Pinto and Prescott [28] are some examples of those.

2.2. Project Success/Failure Factors. One of the most relevant fields of study in project management is success factors and failure causes in projects. They were first identified at the PMI Annual Seminar and Symposium in 1986 [29] and became one of the most discussed themes within specialized literature [30]. Each stakeholder working on a project has his/her own opinion on what is determinant for failure and it is much more complex than adhering to the traditional criteria of time, cost, and quality. If a project can be considered failed when it has not delivered what was expected after its completion, causes leading to this unsatisfactory result are also subjected to the different points of view of the stakeholders involved.

The different lists of success factors and failure causes in the literature show no consensus. The most usual would be a combination of several factors, with different levels of influence in different stages of the project's life cycle, resulting in its success or failure. The concepts of success factors and failure causes are closely related, but a failure cause is not necessarily the negation of, or the opposite to, a success factor. There is not always such correspondence among them. This study is considering failure causes rather than successful factors because we are taking the assumption that the more the failure causes concur in a project, the higher the complexity is.

The literature identifying success/failure factors is very extensive. Some examples, especially regarding failure causes, are [31–35], all of them related to the construction sector. Another sector with many references is the IT (information technology) projects. Some examples are [36–39]. There are several frameworks classifying the factors. Belassi and Tukel [40] suggested a scheme that classifies the critical factors in four different dimensions (Table 1) and describes the impacts of these factors on project performance. Shenhar et al. [25] have identified also four distinct dimensions: project efficiency, impact on the customer, direct and business success, and preparing for the future. They stated that the exact content of each dimension and its relative importance may change with time and are contingent on the specific

Table 1: Belassi and Tukel taxonomy.

Groups of factors
Related to project
Related to the project managers and the team members
Factors related to the organization
Factors related to the external environment

stakeholder. Lim and Mohamed [41] consider two viewpoints of project success: macro and micro. Regarding the micro, they have identified technical, commercial, finance, risk, environmental, and human related factors.

Richardson [42] and King [43] point out that none of the key success factors described in the literature is responsible, on its own, for ensuring a project's success. They are all interdependent and require a holistic approach. Groups of success factors and their interactions are of prime importance in determining a project's success or failure [44]. Multivariate statistics methods may be very useful with this purpose. Some examples applied in the project management complexity are [4, 45, 46] as well as in the field of project success [47–49]. Clustering methods have been also used in the context of exploring complex relationships within the field of project management, for example, the works [50–52]. Our study uses self-organizing maps (SOM) and clustering techniques to find patterns in a data set of answers coming from a survey.

2.3. Self-Organizing Maps and Applications in Project Management. SOM is an unsupervised neural network proposed by Kohonen [53] for visual cluster analysis. The neurons of the map are located on a regular grid embedded in a low (usually 2 or 3) dimensional space and associated with the cluster prototypes by the connected weights. In the course of the learning process, the neurons compete with each other through the best-matching principle in such a way that the input is projected to the nearest neuron given a defined distance metric. The winner neuron and its neighbors on the map are then adjusted towards the input in proportion with the neighborhood distance; consequently, the neighboring neurons likely represent similar patterns of the input data space. Due to the data clustering and spatialization through the topology preserving projection, SOM is widely used in the context of visual clustering applications. Despite the unsupervised nature, the applicability of SOM is extended to classification tasks by means of a variety of ways, such as neuron labeling method, semisupervised learning, or supervised learning vector quantization (LVQ) [54].

SOM is recognized as a useful technique to analyze high-dimensional data sets and understand their hidden relations. It can be used to manage complexity in large data sets [39]. Nevertheless, there are few cases described in the field of project management. Balsera et al. [55] have exposed the application of SOM to analyze information related to effort estimation and software projects features. MacDonell [56] has reported other multidimensional data study visualization based on SOM, identifying groups of data for similar projects and finding nonlinear relationships within the explored variables. Naiem et al. [57] have used SOM for visualizing the set of candidate portfolios. After reviewing the literature, we can conclude that there is not any former application of SOM for analyzing patterns of project failure causes.

Table 2: Failure causes considered in the study.

Cause	Code
Irruption of competitors	C1
Continuous or dramatic changes to the initial requirements	C2
Customer's requirements inaccurate, incomplete or not defined	C3
Disagreement or conflict of interest among departments	C4
Inaccurate cost estimations	C5
Inaccurate time estimations	C6
Deficient management of suppliers and procurement	C7
Lack of Management support	C8
Lack of previous identification of legislation	C9
Badly defined specifications	C10
Political, social, economic or legal changes	C11
Project manager Lack of commitment	C12
Project manager Lack of communication skills	C13
Project manager Lack of competence	C14
Project manager Lack of vision	C15
Project requirements deficiently documented	C16
Project staff changes	C17
Project team lack of competence	C18
Project team misunderstanding related to customer/user needs	C19
Project team lack of commitment	C20
Public opinion opposition to Project	C21
Quality checks badly performed or not performed at all	C22
Extremely new or complex technology	C23
Unexpected events with no effective response possible	C24
Unrealistic customer expectations	C25
Wrong number of people assigned to the project	C26

3. Research Methodology

As discussed in the Introduction, the ground for this work is the questionnaire that was designed to gather information on the perception project managers have of what the success factors and failure causes are. After the information was gathered, a multivariate analysis was performed on the data with cluster data mining techniques.

The sections of the questionnaire considered for this study were as follows:

(i) General information on the respondent and typology of projects he/she was involved in: country, type, and size of project.

(ii) Frequency of different causes of project failure, with 26 multiple-choice questions (from 0–25%, rare or improbable occurrence, to 75–100%, always occurs).

The causes of failure are extracted from the existing bibliography and the previous work on the matter. Table 2

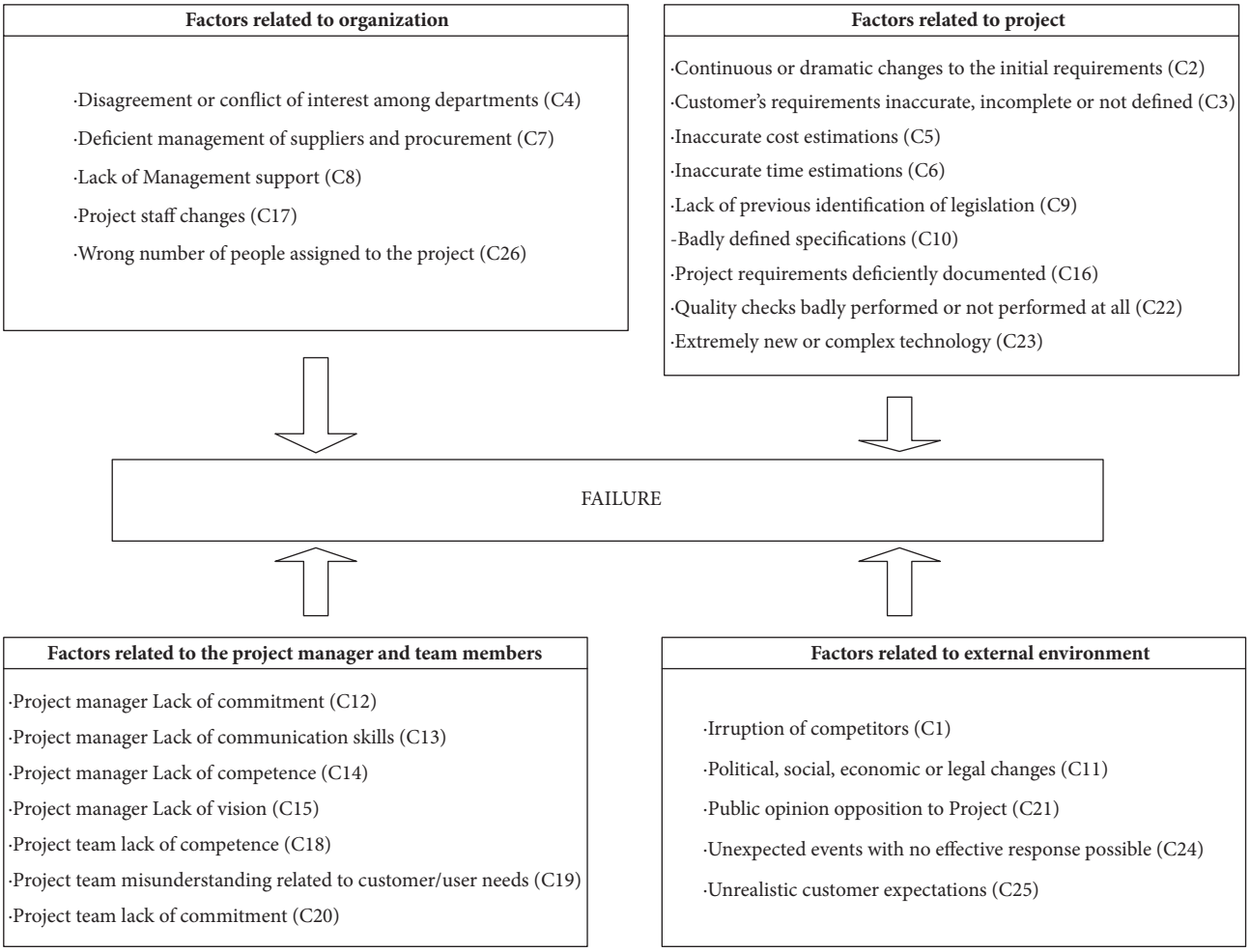

Figure 1: Failure causes grouped by category. The factors are derived from existing bibliography and previous work on the matter. They have been classified according to Belassi and Tukel taxonomy.

gathers the identified factors. The factors are coded as Cx, where x is the number given to the failure cause.

The factors were classified according to Belassi and Tukel taxonomy [40] included in Table 1. Figure 1 depicts the classification considered. The factors are presented unshorted because they were shuffled in order to avoid biases in the survey. The first group includes the factors related to the organization the project belongs to (i.e., factors related to top management support). The second group includes the factors related to the project itself and the way the project is managed. The third one comprises the factors related to the project manager and the team members. Many studies demonstrated the importance of selecting project managers who possess the necessary technical and administrative skills for successful project termination, as well as their commitment. The competences of the team members are also found to be a critical factor. Finally, the factors related to external environment consist of factors which are external to the organization but still have an impact on project success or failure. The classification can be considered collectively exhaustive. Belassi and Tukel state that the four groups offer a comprehensive set in that any factor listed in the literature, or even specific points of consideration, should belong to at least one group [40]. However, it is not easy to differentiate in which category to include each of the different factors. The groups are interrelated. The authors give some indications to distinguish where to classify them, mainly regarding organization and external environment. For example, if a customer is from outside the organization, he/she should be considered as an external factor (i.e., C25-unrealistic customer expectations). For functional projects, however, customers are usually part of the organization, such as top management. In such cases, factors related to the client can be grouped under the factors related to the organization.

We can find correspondences among most of the failure causes included in our study (Table 2) and the complexity indicators appointed by Kermanshachi et al. [3]. For example, C2 (which is one of the most frequent in accordance with the rank presented in Table 7) is related to several complexity indexes as "level of project design changes derived by RFI" and "magnitude of change orders". C10, for example, is related to "percentage of design completed at the start of construction". This fact reinforces the connection between failure causes and complexity factors.

The recipients were randomly chosen among the members of 36 project management groups from LinkedIn network. The questionnaire was open for 3 months and 11 days in 2011, in order to obtain a significant number of answers. During that period of time, customized emails were sent with an invitation to answer the questionnaire to a total of 3,668 people. A total of 619 answers were received (16.88%), 611 of which were considered for further analysis (the rest were discarded due to consistency issues). Neither of the questionnaires was partially filled or incomplete, since all fields were marked as mandatory. Previously, in 2010, a pilot survey was conducted with the help of project management experts and practitioners. This primary questionnaire was sent up to 45 people with a response rate of 66.67%. Those factors which reached a higher score were the ones finally selected to configure the list included in the current study. Suggestions and comments made by respondents to improve the proposed list were also taken into account.

Answers from 63 countries were received, as Table 3 shows.

The answers have been grouped first into 13 different geographical zones (more details can be found in [16]), according to geographical, cultural, historical, and economic criteria. From those, a total of 6 groups comprise more than 85% of the respondents: AMZ3, EUZ3, EUZ1, AMZ1, ASZ1, and ASZ3. The other 15% of the responses were grouped in a new category called "Others". Table 4 summarizes the groups by geographical zone.

Following a similar approach, 53 types of projects were present on the answers received. To simplify, they have been classified into a total of 17 groups, according to the highest level of the ISIC Rev. 4 Classification [58]. Development aid projects were listed apart as an independent category. Table 5 presents the number of respondents from each project type and the codification.

Information about the size of the projects the respondents are usually involved in was also enquired. The size of the projects is something quite subjective and difficult to compare among different sectors. For example, considering the project budget as an indicator of project size, a project considered small in construction could, by contrast, be considered large in information technology projects. In order to avoid these biases, the guideline provided by the University of Wisconsin-Madison [59] was used. The results show that 11.95% of respondents are involved in small projects, 49.26% in medium size projects, and 38.79% in large projects.

Though the scope of this paper does not cover the statistical analysis of the obtained answers, here is a summary of the main results. Overall, both the most and least frequent failure causes are presented here. In order to rank each one of the 26 failure causes, a frequency index (FI) was calculated as follows:

$$\mathrm{FI} = \frac{\left[\sum_{i=1}^{611} \sum_{j=1}^{4} A_i * B_j\right]}{611}, \tag{1}$$

where

A_i is the number of responses choosing interval i,

B_j is interval factor j.

Table 3: Countries and number of responses (ordered by number of responses). When several countries are indicated in the same row, the number of respondents stands for the number of answers received from each country (i.e., 50 answers from Spain and 50 answers from the United States).

COUNTRIES	NUMBER OF RESPONDENTS
Argentina	91
Spain; United States	50
Greece	45
Chile	41
India	40
Brazil	30
Luxembourg	27
Mexico	21
Uruguay	20
United Arab Emirates	19
Italy	18
Russia	16
Sweden	14
South Africa	12
Canada	11
Germany	8
Saudi Arabia; The Netherlands; United Kingdom	7
Australia; Denmark; Finland; Kuwait	5
Belgium; Paraguay; Slovenia	4
Egypt; France; Israel; Switzerland	3
Philippines; Tanzania	2
Algeria; Austria; Bahrain; Belarus; Botswana; British Virgin Islands; Cameroon; China; Colombia; Cyprus; Czech Republic; Ghana; Hong Kong; Hungary; Indonesia; Ireland; Japan; Jordan; Nigeria; Pakistan; Peru; Poland; Qatar; Rwanda; Singapore; Surinam; Thailand; Uganda; Venezuela	1

Each interval factor is defined as presented in Table 6. The ranking is presented in Table 7.

With the obtained results, a cluster analysis was conducted to obtain patterns in the answers data set, grouping a set of data with similar values. The aim is to classify a set of simple elements into a number of groups in such a way that the elements in the same group are similar or related to one another and, at the same time, different or unrelated to the elements in other groups.

The cluster analysis (also called Unsupervised Classification, Exploratory Data Analysis, Clustering, Numerical Taxonomy, or Pattern Recognition) is a multivariate statistical technique whose aim is to divide a set of objects into groups or *clusters* in such a way that objects in the same cluster are very similar to each other (internal cluster cohesion) and the objects in other groups are different (external cluster isolation). Summarizing, it deals with creating data clusters

Table 4: Answers grouped by geographical zone.

GEOGRAPHICAL ZONES ANALYZED	COUNTRIES	NUMBER OF RESPONDENTS
AMZ3	Argentina, Brazil, Chile, Colombia, Paraguay, Peru, Surinam, Uruguay and Venezuela	190
EUZ3	Cyprus, Spain, Greece and Italy	114
EUZ1	Germany, Austria, Belgium, Denmark, Finland, France, Ireland, Virgin Islands (UK), Luxembourg, United Kingdom, The Netherlands, Sweden and Switzerland	86
AMZ1	Canada and United States	61
ASZ1	Algeria, Bahrain, Indonesia, Jordan, Pakistan, Qatar, Saudi Arabia, Egypt, United Arab Emirates and Kuwait	40
ASZ3	India	40
OTHERS	Rest	80

Table 5: Project types and number of responses.

CODE	PROJECT TYPE	NUMBER OF RESPONDENTS
ISIC1	Information and communication	365
ISIC2	Financial and insurance activities	43
ISIC3	Construction	42
ISIC4	Manufacturing	41
ISIC5	Professional, scientific and technical activities	32
ISIC6	Public administration and defence	12
ISIC7	Mining and quarrying	11
ISIC8	Human health and social work activities	10
ISIC9	Electricity, gas, steam and air conditioning supply	10
ISIC10	Accommodation and food service activities	8
ISIC11	Education	8
ISIC12	Administrative and support service activities	7
ISIC13	Transportation and storage	7
ISIC14	Wholesale and retail trade	6
ISIC15	Real estate activities	4
ISIC16	Development aid	3
ISIC17	Arts, entertainment and recreation	2

Table 6: Intervals and their interval factor.

Interval	Meaning	Factor (interval arithmetic mean)
0–25%	Seldom/Unlikely to occur	0,125
25–50%	Occasional/Likely to occur	0,375
50–75%	Frequent/Very likely to occur	0,625
75–100%	Always occurs	0,875

in such a way that each group is homogeneous and different from the rest. For this purpose, many data analysis techniques can be used. In this study, the Self-organized Maps (SOM) technique has been used, a specific type of neural network [60].

SOM networks are an excellent tool for exploring and analyzing data, which are especially adequate due to their remarkable visualization properties. They create a series of prototype vectors which represent the data set and project such vectors into a low dimensional network (generally bidimensional) from the input d-dimensional space, which preserves its topology and maintains it. As a result, the network shows the distance between the different sets, so that it can be used as an adequate visualization surface to display different data characteristics as, for instance, their cluster division. Summarizing, SOM Clustering Networks allow input data clustering and easily visualize the resulting multidimensional data clusters. For the analysis of data collected in the questionnaire, SOM Toolbox was used with MATLAB [61]. The methodology used is a two-level approach for partitive clustering, where the data set is first projected using the SOM, and then the SOM is clustered, as described in [62]. Partitive clustering algorithms divide a data set into a number of clusters, typically by trying to minimize some criterion or error function. An example of a commonly used partitive algorithm is the k-means, which minimizes error function (2), where C is the number of clusters and c_k is the center of cluster k.

$$E = \sum_{k=1}^{C} \sum_{x \in Q_k} \left\| x - c_k \right\|^2 \tag{2}$$

Table 7: Summary of results.

RANK	CODE	FAILURE CAUSE	FREQUENCY INDEX
1	C3	Customer's requirements inaccurate, incomplete or not defined	0,5992
2	C2	Continuous or dramatic changes to initial requirements	0,5665
3	C6	Inaccurate time estimations	0,5329
4	C16	Project requirements inadequately documented	0,5092
5	C10	Not or badly defined specifications at the time the Project Team starts to work	0,4859
6	C5	Inaccurate costs estimations	0,4826
7	C8	Lack of Management support	0,4605
8	C25	Unrealistic Customer's expectations	0,4507
9	C4	Disagreements or conflicts of interest among different departments involved	0,4478
10	C19	Project Team's misunderstandings related to Customer/User's wishes or needs	0,4417
11	C17	Project staff changes	0,4306
12	C22	Quality checks not or badly performed	0,4270
13	C26	Wrong number of people assigned to the project (too many, too few)	0,4135
14	C13	Project Manager's lack of communication skills	0,4114
15	C7	Inadequate management of suppliers and procurement	0,3828
16	C18	Project Team's lack of competence	0,3815
17	C15	Project Manager's lack of vision	0,3746
18	C14	Project Manager's lack of competence	0,3742
19	C23	Too much complex or new technology	0,3574
20	C20	Project Team's lack of commitment	0,3537
21	C24	Unexpected events with no effective response possible	0,3292
22	C9	Lack of previous identification of relevant rules and legislation	0,3194
23	C12	Project Manager's lack of commitment	0,3009
24	C11	Political, social, economic or legal changes	0,2662
25	C21	Public opinion opposition to project	0,2428
26	C1	Competitors	0,2261

To select the best one among different partitionings, each of these can be evaluated using some kind of validity index. Several indices have been proposed [63, 64]. In our work, we used the Davies-Bouldin index [65], which has been proven to be suitable for evaluation of k-means partitioning. This index is a function of the ratio of the sum of within-cluster scatter to between-cluster separation. According to Davies-Bouldin validity index, the best clustering minimizes (3), which uses S_c for within-cluster distance ($S(Q_k)$) and d_{ce} for between-cluster distance ($d(Q_k, Q_l)$).

$$\frac{1}{C}\sum_{k=1}^{C}\max_{l\neq k}\left\{\frac{S_c(Q_k)+S_c(Q_l)}{d_{ce}(Q_k,Q_l)}\right\} \tag{3}$$

The approach used in this paper (clustering the SOM rather than clustering the data directly) is depicted in Figure 2. First, a large set of prototypes (much larger than the expected number of clusters) is formed using SOM. The prototypes can be interpreted as "protoclusters" [62], which are in the next step combined to form the actual clusters. Each data vector of the original data set belongs to the same cluster as its nearest prototype.

The SOM consists of a regular, usually two-dimensional, grid of map units. Each unit i is represented by a prototype vector $m_i = [m_{i1}, \ldots, m_{id}]$, where d is input vector dimension. The units are connected to adjacent ones by a neighborhood relation. The number of map units, which typically varies from a few dozen up to several thousand, determines the accuracy and generalization capability of the SOM. During training, the SOM forms an elastic net that folds onto the "cloud" formed by the input data. Data points lying near each other in the input space are mapped onto nearby map units. Thus, the SOM can be interpreted as a topology preserving mapping from input space onto 2D grid of map units.

The SOM is trained iteratively. At each training step, a sample vector x is randomly chosen from the input data set. Distances between x and all the prototype vectors are computed. The best-matching unit (BMU), which is denoted here by b, is the map unit with prototype closest to x.

$$\|x - m_b\| = \min_i \{\|x - m_i\|\} \tag{4}$$

Next, the prototype vectors are updated. The BMU and its topological neighbors are moved closer to the input vector in

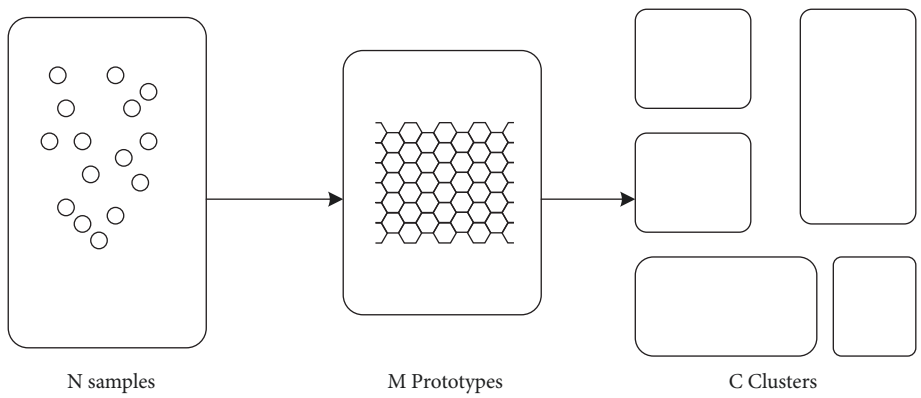

Figure 2: Clustering process.

the input space. The update rule for the prototype vector of unit i is

$$m_i(t+1) = m_i(t) + \alpha(t)\, h_{bi}(t)\left[x - m_i(t)\right], \tag{5}$$

where

t is time,

$\alpha(t)$ is the adaptation coefficient,

$h_{bi}(t)$ is neighborhood kernel centered on the winner unit:

$$h_{bi}(t) = \exp\left(-\frac{\left\|r_b - r_i\right\|^2}{2\sigma^2(t)}\right), \tag{6}$$

where r_b and r_i are positions of neurons b and i on the SOM grid. Both $\alpha(t)$ and $\sigma(t)$ decrease monotonically with time.

To perform the analysis, a file with the following 30 input variables was prepared (the columns of the data matrix are the variables and each row is a sample):

(i) Project size (this categorical variable has been encoded as follows: small = 3; medium = 3^2 = 9; large = 3^3 = 27).

(ii) 26 failure causes (Table 2).

(iii) Answer ID (encoded from P1 to P611). This variable is included only for tracing purpose during the validation stage, not for training.

(iv) Country (the countries were grouped into 13 geographical areas, taking into account geographical, economic, historical, and cultural criteria).

(v) Type of project (they are encoded into 17 activities derived from the ISIC/CIIU Rev. 4 codes).

For the training, only the scores of the 26 failure causes and the project size were used. The other variables of the data set (country and type of project) were used just only in analyzing the results of the clustering. The training has been performed trying different grid dimensions but always considering hexagonal topology (6 adjacent neighbors) and trying different numbers of clusters. One of the most significant pieces of information provided by this kind of analysis is precisely the optimal number of clusters or *optimal clustering* and how the samples are distributed among the clusters. In order to determine the right number of clusters, the Davies-Bouldin index (DBI) was used [65] as described above.

The resulting clusters were analyzed in order to characterize the perception of project manager practitioners about project failure and hence complexity. An analysis of which factors make each cluster different from the rest of the data is performed, as well as what the dependencies between variables are in the clusters. Each cluster was featured finding the set of failure causes rated over or under the mode for the global survey and identifying the most remarkable differences with the other clusters. In order to do that, histograms and radar charts were used. Finally, we studied the population distribution of each cluster regarding geographical areas, projects sector, and projects size. The analysis is presented by means of contingency tables.

4. Results

After several preliminary trials, a 7×5 hexagonal SOM topology was chosen. The SOM was trained following the method described in the former section. Figure 3 shows the Unified distance matrix (U-matrix) and 27 component planes, one for each variable included in the training. The U-matrix is a representation where the Euclidean distance between the codebook vectors of neighboring neurons are depicted in a color schema image. In this case, high values are associated with red colors and low values with blue colors. This image is used to visualize the data in a high-dimensional space using a 2D image. High values in the U-matrix represent a frontier region between clusters, and low values represent a high degree of similarity among neurons on that region. Each component plane shows the values of one variable in each map unit. Through these component planes, we can realize emerging patterns of data distribution on SOM's grid and detect correlations among variables and the contribution of each one to the SOM differentiation only viewing the colored pattern for each component plane [66].

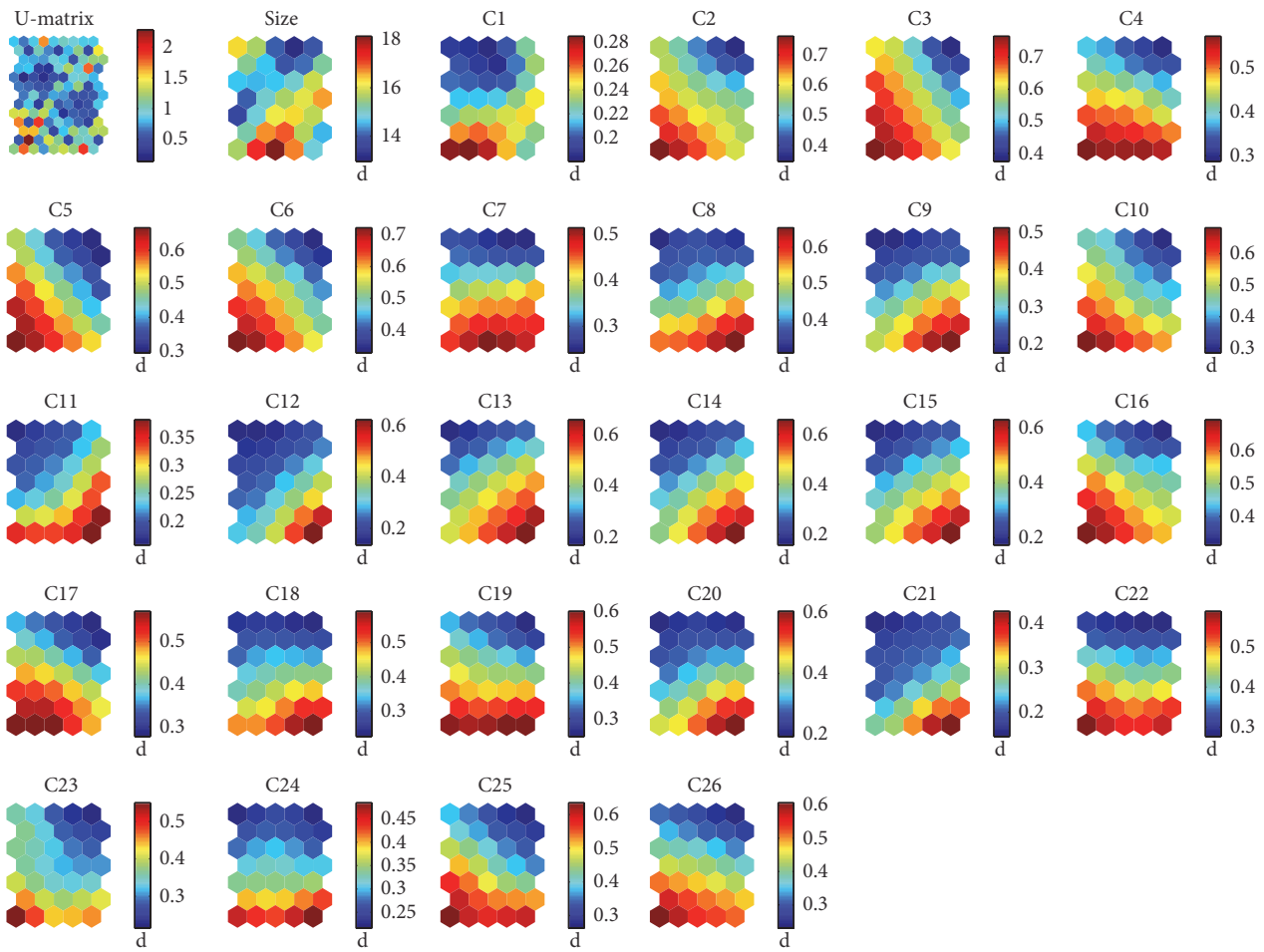

Figure 3: SOM U-matrix and component plane matrices using a 7 × 5 hexagonal SOM.

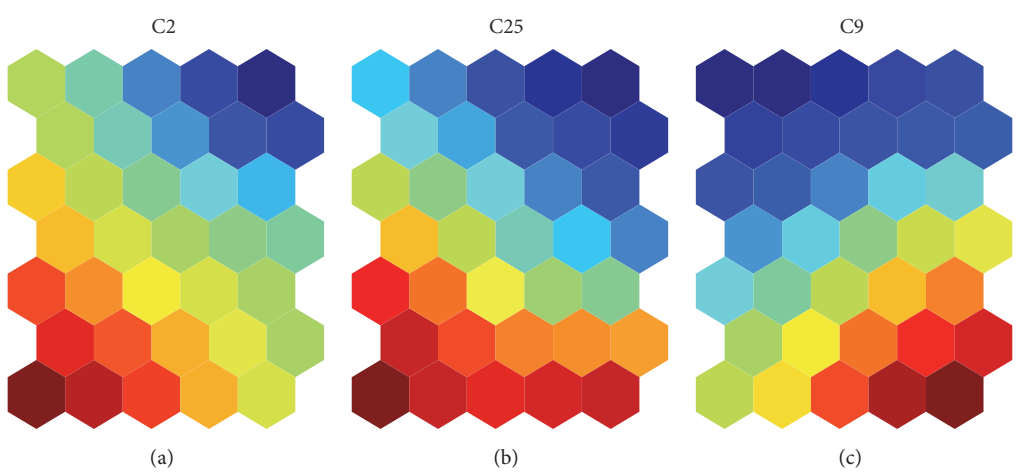

Figure 4: Prototypes of the three patterns found in the component planes.

Several similarities among the component planes of some factors can be found in Figure 3. Three different patterns can be clearly distinguished:

(i) Pattern I: C2, C3, C5, C6, C10, C16, and C17 (Figure 4(a))

(ii) Pattern II: C25 and C26 (Figure 4(b))

(iii) Pattern III: C9, C12, C13, C14, C15, C18, C20, and C21 (Figure 4(c))

An example of each component plane has been depicted in Figure 4 to illustrate each of the patterns. As introduced

Table 8: Summary of patterns interpretation.

Pattern number	Focus attention	Interpretation
Pattern I	How the project is managed	This pattern of respondents links project complexity mainly with the characteristics of the project and how it is managed: the requirements are incomplete or inaccurate, with continuous changes, the specifications are badly defined and both costs and schedule are inaccurate. They consider also that the project staff changes entail a higher complexity.
Pattern II	Insufficient number of resources and unrealistic customer expectations	They emphasize the influence of unrealistic customer expectations and a wrong (it is supposed to be insufficient) number of people assigned to the project. It is curious that neither of them is apparently under their responsibility (under the hypothesis that the resources assigned to the project are a decision made by the organization managers).
Pattern III	Project manager and team members skills, competences and commitment	They have a more personalistic view of the complexity, with the role of the project manager and team members being decisive.

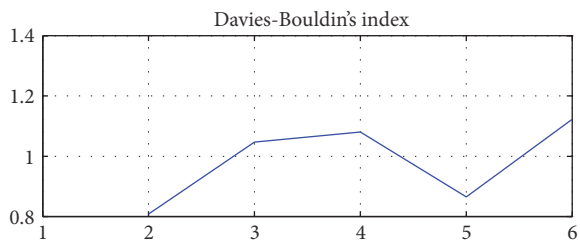

Figure 5: Davies-Bouldin's index using a hexagonal 7 × 5 SOM calculated for each clustering (x-axis: number of clusters/y-axis: Davies-Bouldin's index).

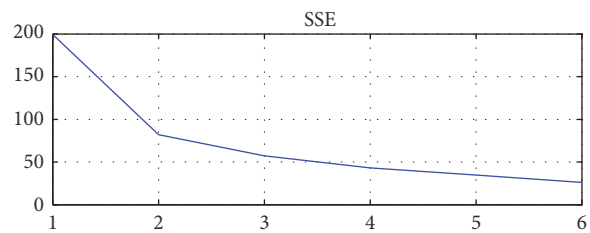

Figure 6: Sum of squared errors using a hexagonal 7 × 5 SOM calculated for each clustering (x-axis: number of clusters/y-axis: sums of squared errors).

previously, we can detect correlations among variables by just looking at similarities in the component planes [66]. We can infer combinations of answers from the surveyed project managers. In this case, a combination of answering patterns can be concluded from each of the three groups. So, the project managers that rate factor C2 as frequent also rate as frequent factors C3, C5, C6, C10, C16, and C17 *(Pattern I)*. It is also the same with the other two patterns. Regarding *Pattern I*, it should be noticed that all the factors considered belong to the project category, except factor C17 that belongs to the organization category. Regarding *Pattern II*, there is an association between the rating of unrealistic customer's expectations (C25) and a wrong number of people assigned to the project (C26). Regarding *Pattern III*, most of the factors belong to the project manager and team members (noted in the figures as PM&T) category (C12, C13, C14, C15, C18, and C20).

Table 8 summarizes the interpretation of each pattern. A focus of attention has been remarked for each one and how each pattern could be understood. Obviously, interpretation has a component of subjectivity.

Next, k-means algorithm was used to build the clusters and the Davies-Bouldin index was considered to determine the optimum number of clusters. The Davies-Bouldin index and the sums of squared errors diagrams are displayed in Figures 5 and 6, where the x-axis represents the number of clusters. According to the usual work methodology of this type of neural networks, the best possible clusterization will be the one that reaches a better compromise in the minimization of both parameters. In this case, we have taken 5 different clusters.

Figure 7 depicts the clustering results taking as reference the results of the application of the k-means algorithm on the optimal cluster number according to Davies-Bouldin. The SOM plot sample hits (b) represent the number of input vectors classified by each neuron. The relative number of vectors for each neuron is shown via the hexagon's size and its color represents the similarity among neurons.

Pearson's chi-squared test [67] was calculated to find dependence among the variables and their distribution in the clusters. The resulting p values for the contingency tables shown in Tables 9, 10, and 11 were higher than 0.05, so it can be concluded that neither the project size nor the geographical zone or the project type is significant for the groups determined. Nevertheless, the p values obtained for Cx are less than 0.05 which implies that the variation of the values of Cx through the clusters has statistical significance, so we proceeded to the study and categorization of clusters based on these factors with statistical significance. Bar charts were plotted to characterize each cluster (depicted in Figures 8, 9, 10, and 11). Factors have been grouped by their category according to Belassi and Tukel taxonomy [40] in order to facilitate the interpretation of the information.

5. Discussion of Results

It can be observed that clusters 3 and 5 are associated with the highest rates of the factors for each category, while clusters

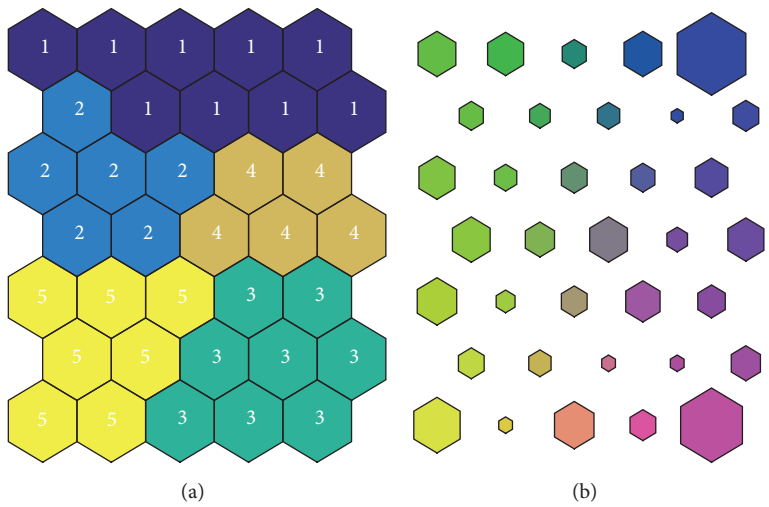

Figure 7: Partitioning of the SOM codebooks with 5 clusters (a) and the SOM plot sample hits (b).

Table 9: Contingency table showing the distribution of project size among the clusters.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Samples
Size						
Small	25	14	14	6	14	73
Medium	84	53	68	40	56	301
Large	57	35	61	42	42	237
Samples	166	102	143	88	112	611

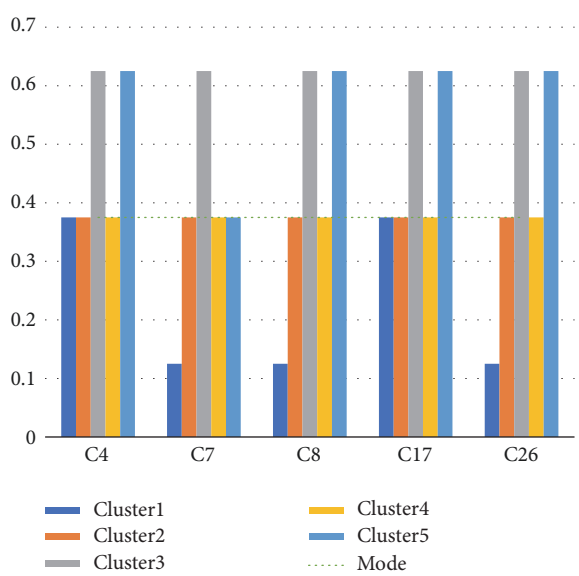

Figure 8: Organization factors and clusters bar chart. Factor importance average score of each cluster plotted against the mode of the global survey.

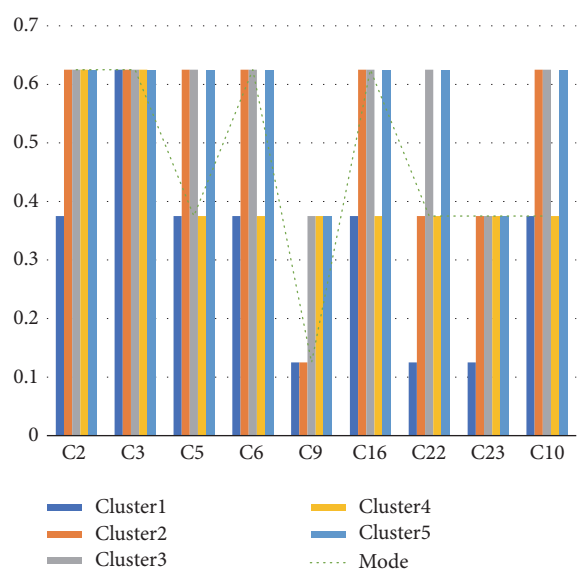

Figure 9: Project factors and clusters bar chart. Factor importance average score of each cluster plotted against the mode of the global survey.

1 and 2 are the lowest. In fact, cluster 1 only attributes high importance/frequency to C3 (customer's requirements inaccurate, incomplete, or not defined). On the other hand, cluster 3 rates with high values 19 in 26 factors, and cluster 5 rates high values 13 in 26. So, the general trend is that clusters 3 and 5 rate factors equal to or over the general mode and clusters 1 and 2 rate factors equal to or below the general mode, as represented in Figures 8, 9, 10, and 11. An exception of this can be found for C5 (inaccurate cost estimations), where cluster 2 rates higher than the mode. A global rating

Table 10: Contingency table showing the distribution of geographical zone among the clusters (see Table 4 for details about the geographical zones).

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Samples
Zone						
AMZ1	26	9	11	5	10	61
AMZ3	43	35	48	30	34	190
ASZ1	8	3	12	9	8	40
ASZ3	9	5	13	7	6	40
EUZ1	23	18	17	10	18	86
EUZ3	38	19	22	16	19	114
Others	19	13	20	11	17	80
Samples	166	102	143	88	112	611

Table 11: Contingency table showing the distribution of project type among the clusters (see Table 5 for details about the project type codes).

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Samples
Type						
ISIC1	100	59	90	49	67	365
ISIC2	14	5	9	3	12	43
ISIC3	12	5	10	8	7	42
ISIC4	7	7	12	8	7	41
ISIC5	10	9	3	6	4	32
ISIC6	1	0	5	4	2	12
ISIC7	1	1	4	3	2	11
ISIC8	3	3	3	0	1	10
ISIC9	6	1	2	1	0	10
ISIC10	1	2	1	1	3	8
ISIC11	2	2	1	0	3	8
ISIC12	3	3	0	1	0	7
ISIC13	2	2	0	1	2	7
ISIC14	3	2	0	1	0	6
ISIC15	0	0	1	1	2	4
ISIC16	0	0	2	1	0	3
ISIC17	1	1	0	0	0	2
Samples	166	102	143	88	112	611

overview of each cluster can be drawn from Figure 12, where a radar chart was built representing the average rates of each cluster to the factors included in each category.

In a similar way, from Figure 13, we can conclude how each group of factors is rated within each cluster. It can be observed that all the factors are arranged following the same order (project, organization, project manager and team members, and external factors) with three exceptions: cluster 1 gives PM&T the same importance as the external factors (very low in both cases), cluster 4 gives PM&T the same importance as organization factors, and cluster 3 is the most remarkable because it gives the same importance to the project, the organization, and the PM&T factors.

An interpretation of each cluster (Table 12) can be inferred from the presented results, especially Figure 13. In order to facilitate the understanding, the focus of attention of each cluster is remarked.

Finally, we can also find some connections between the patterns found in the component planes (summarized in Table 8) and the clusters. The most remarkable connections can be found within *Pattern I* and Cluster 1 (both sharing a high importance of incomplete or inaccurate requirements, the badly defined specifications, the project staff changes, and inaccurate costs estimations) and *Pattern III* and Cluster 3 (both remark the importance of project management and team members skills, competences, and commitment).

6. Conclusions

The analysis performed by clustering techniques has allowed us to conclude that the total number of answers obtained can be grouped into 5 classes of respondents, who behaved differently in analyzing project failure causes and hence project complexity. This result is coherent with the conclusions of the existing literature on the subject, which claims that there are no unique criteria and subjectivity is an inherent characteristic of these assessments. The behavior of each cluster can be understood by means of the bar charts shown.

Table 12: Summary of clusters interpretation.

Cluster number	Focus attention	Description
Cluster 1	Project factors, specially customer's requirements inaccurate, incomplete or not defined.	Cluster 1 is featured as the pattern of respondents that gives very low influence over the projects failure to the factors included in the questionnaire. They just only attribute some influence to the project and the organization categories. The most remarked factors are (highly rated factors that are equal or above the general mode): (i) C4: Disagreement of conflict of interest among departments (ii) C17: Project staff changes (iii) C3: Customer's requirements inaccurate, incomplete or not defined (iv) C5: Inaccurate cost estimations (v) C10: Badly defined specifications
Cluster 2	Both the project and the organization, remarking the influence of inaccurate cost estimations.	Cluster 2 attributes the highest influence to the project category, followed by the factors related to the organization and, to a lesser extent, to the project manager and, team members and external categories. All the factors related to the organization are scored with the same importance as the general mode. Regarding the project category, all the factors are rated high or medium, excepting C9 (lack of previous identification of legislation). The stress is specially put in C5 (inaccurate cost estimations).
Cluster 3	Project, organization and, project manager and team members factors are all considered at the highest value.	Cluster 3 is featured as the pattern of respondents that consider the project manager and team members factors as the highest value, at the same level as the project and the organization factors.
Cluster 4	Project factors mainly, but organization and project manager and team members factors are also very important.	Cluster 4 is characterized because it associates a medium influence to project, organization, and project manager and team member almost equally. All the factors are rated equal to the mode, except three: (i) C6: Inaccurate time estimations (rated below the mode) (ii) C9: Lack of previous identification of legislation (rated above the mode) (iii) C16: Project requirements deficiently documented (rated below the mode)
Cluster 5	Project and organization	Cluster 5 associates a very high influence on the projects failure to the categories of project and organization. The main difference with cluster 3 is that cluster 5 attributes less importance to the project manager and team members category.

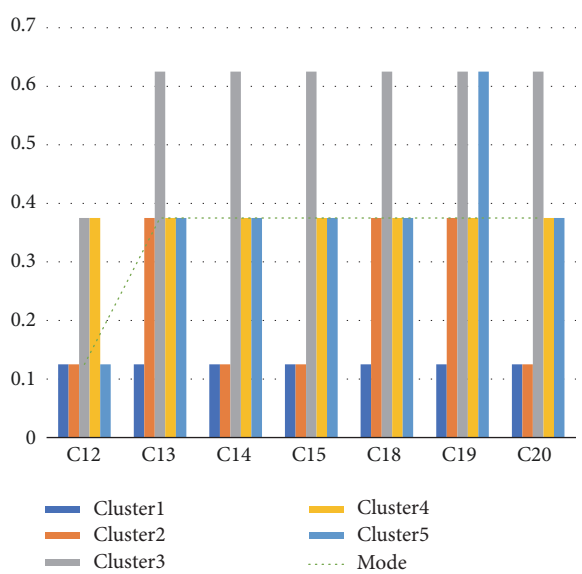

Figure 10: Project manager and team members factors bar chart. Factor importance average score of each cluster plotted against the mode of the global survey.

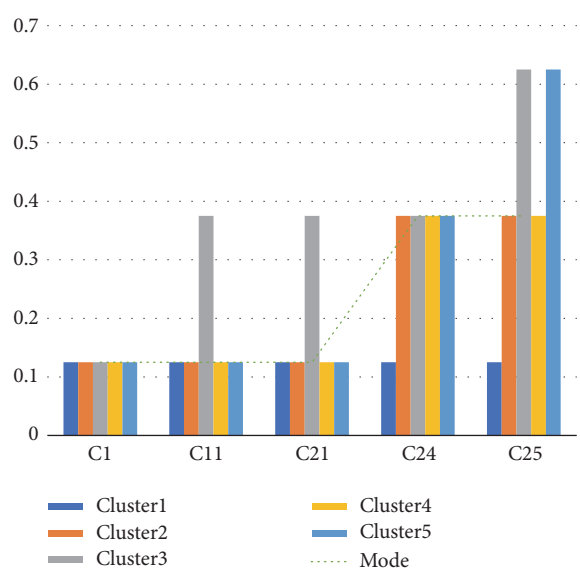

Figure 11: External factors and clusters bar chart. Factor importance average score of each cluster plotted against the mode of the global survey.

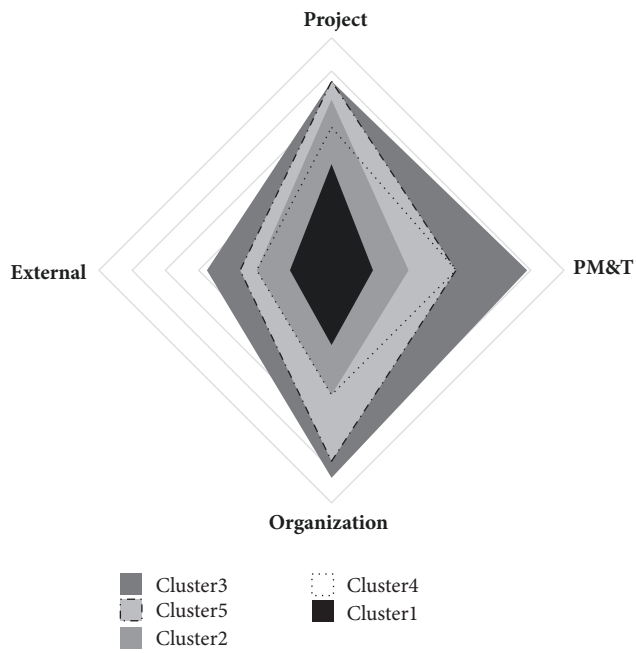

Figure 12: Radar chart representing the average rate for the factors of each category.

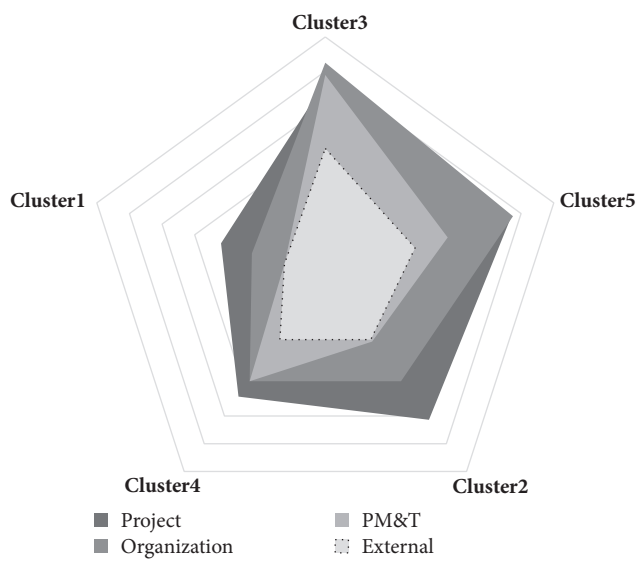

Figure 13: Radar chart representing the average rate of each cluster in each category.

Representing the average of each cluster for each variable and comparing it to the global mode have proven to be meaningful in characterizing each cluster.

Regarding the clusters identified and based on the distribution of samples and the characterization of each one, we can conclude that they are representative, in the sense that each one has its own differential set of features. It is remarkable that neither the project size nor the geographical area or the project type is significant considering the clusters, so it can be concluded that the answers given are not specific to a particular country or type/size of project.

The prevalent order of importance for the factor categories is project, organization, project manager and team members, and external factors. The most remarkable exception can be found in cluster 3, which attributes the highest influence also to the project manager and team members category. The influence of external factors is in all the cases very low, so it can be remarked that the project managers attribute the complexity to the inner features of the projects and the management conditions.

On the other hand, the analysis of the component planes has revealed that there are three clear patterns. The first one (Pattern I) establishes a close association with the factors included in the project category as factors of failure and hence complexity. However, the third pattern (Pattern III) focuses on the factors associated with the project manager and the team members. This is a remarkable fact that shows at least two schools of thought: the one considering the factors inherent to the project and the one that attributes the complexity to the skills and competences of the project managers and the team members. The second pattern points out a relation between factors C25 and C26, which belong to two different categories. The analysis of the component planes is independent of the cluster analysis, although some similarities can be found between these patterns and the clusters. The most remarkable are the similarities of *Pattern I* and Cluster 1 and *Pattern III* and Cluster 3.

Finally, this study provides a multidimensional analysis of the complexity in projects. Some significative combinations of factors have been found. A limitation of this study is that only the perception of project managers is considered. The study may be extended considering other project stakeholders. Other limitations are that more than 67% of the answers come from only 10 countries (Argentina, Spain, United States, Greece, Chile, India, Brazil, Luxembourg, Mexico, and Uruguay) and there are a number of countries with a very limited number of answers. In addition, more than 59% of the answers are related to IT projects. Taking into account the precedent limitations detected in the study, results are likely to be biased because of country and type of project response percentages, and therefore they cannot be generalized. The results have been exposed per geographical zone and per project type to avoid the referred limitation.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work has been subsidized through the Plan of Science, Technology and Innovation of the Principality of Asturias (Ref. FC-15-GRUPIN14-132).

References

[1] B. Dao, S. Kermanshachi, J. Shane, S. Anderson, and E. Hare, "Identifying and Measuring Project Complexity," in *Proceedings of the International Conference on Sustainable Design, Engineering and Construction, ICSDEC 2016*, pp. 476–482, usa, May 2016.

[2] K. I. Gidado, "Project complexity: The focal point of construction production planning," *Construction Management and Economics*, vol. 14, no. 3, pp. 213–225, 1996.

[3] S. Kermanshachi, B. Dao, J. Shane, and S. Anderson, "An Empirical Study into Identifying Project Complexity Management Strategies," in *Proceedings of the International Conference on Sustainable Design, Engineering and Construction, ICSDEC 2016*, pp. 603–610, usa, May 2016.

[4] L.-A. Vidal and F. Marle, "Understanding project complexity: implications on project management," *Kybernetes*, vol. 37, no. 8, pp. 1094–1110, 2008.

[5] K. Remington, R. Zolin, and R. Turner, "A model of project complexity: distinguishing dimensions of complexity from severity," in *Proceedings of the 9th International Research Network of Project Management Conference*, 2009.

[6] B. N. Baker, D. C. Murphy, D. Fisher, D. I. Cleland, and W. R. King, "Project management handbook," in *Project Management Handbook New York*, pp. 902–919, Van Nostrand Reinhold, 1988.

[7] D. K. Peterson, C. Kim, J. H. Kim, and T. Tamura, "The perceptions of information systems designers from the United States, Japan, and Korea on success and failure factors," *International Journal of Information Management*, vol. 22, no. 6, pp. 421–439, 2002.

[8] T. Huysegoms, M. Snoeck, G. Dedene, A. Goderis, and F. Stumpe, "A case study on variability management in software product lines: Identifying why real-life projects fail," *International Journal of Information Systems and Project Management*, vol. 1, no. 1, pp. 37–48, 2013.

[9] H. Ernst, "Success factors of new product development: A review of the empirical literature," *International Journal of Management Reviews*, vol. 4, no. 1, pp. 1–40, 2002.

[10] V. R. Montequín, S. C. Fernández, F. O. Fernández, and J. V. Balsera, "Analysis of the Success Factors and Failure Causes in Projects," *International Journal of Information Technology Project Management*, vol. 7, no. 1, pp. 18–31, 2016.

[11] R. Osei-Kyei and A. P. C. Chan, "Review of studies on the Critical Success Factors for Public-Private Partnership (PPP) projects from 1990 to 2013," *International Journal of Project Management*, vol. 33, no. 6, pp. 1335–1346, 2015.

[12] J. R. Turner and R. Müller, *The project manager's leadership style as a success factor on projects: A literature review*, Project Management Institute, 2005.

[13] U. Shauchenka, *Why Projects Fail. Uladzislau Shauchenka*, ISBN 0987724207, 2013.

[14] K. G. Cooper, J. M. Lyneis, and B. J. Bryant, "Learning to learn, from past to future," *International Journal of Project Management*, vol. 20, no. 3, pp. 213–219, 2002.

[15] T. Williams, "Assessing and moving on from the dominant project management discourse in the light of project overruns," *IEEE Transactions on Engineering Management*, vol. 52, no. 4, pp. 497–508, 2005.

[16] S. M. Cousillas Fernández, *Valoración y cuantificación de causas de fracaso y factores de éxito en proyectos*, Universidad de Oviedo, 2013.

[17] S. M. Cousillas, V. Rodríguez, R. Concepción, and F. Rodríguez, "Dentification and analysis of causes for project failure and factors for project success in the asturian case," in *Proceedings of the XIV International Congress on Project Engineering*, 2010.

[18] D. Baccarini, "The concept of project complexity—a review," *International Journal of Project Management*, vol. 14, no. 4, pp. 201–204, 1996.

[19] M. Bosch-Rekveldt, Y. Jongkind, H. Mooi, H. Bakker, and A. Verbraeck, "Grasping project complexity in large engineering projects: The TOE (Technical, Organizational and Environmental) framework," *International Journal of Project Management*, vol. 29, no. 6, pp. 728–739, 2011.

[20] S. Cicmil, T. Cooke-Davies, L. Crawford, and K. Richardson, *Exploring the complexity of projects: Implications of complexity theory for project management practice*, Project Management Institute, 2017.

[21] T. Cooke-Davies, *Aspects of complexity: Managing projects in a complex world*, Project Management Institute, 2011.

[22] C. Ivory and N. Alderman, *Can project management learn anything from studies of failure in complex systems?* Project Management Institute, 2005.

[23] L.-A. Vidal, F. Marle, and J.-C. Bocquet, "Measuring project complexity using the Analytic Hierarchy Process," *International Journal of Project Management*, vol. 29, no. 6, pp. 718–727, 2011.

[24] Y. Lu, L. Luo, H. Wang, Y. Le, and Q. Shi, "Measurement model of project complexity for large-scale projects from task and organization perspective," *International Journal of Project Management*, vol. 33, no. 3, pp. 610–622, 2015.

[25] A. J. Shenhar, O. Levy, and D. Dvir, "Mapping the dimensions of project success," *Project Management Journal*, vol. 28, pp. 5–13, 1997.

[26] D. Dvir and T. Lechler, "Plans are nothing, changing plans is everything: The impact of changes on project success," *Research Policy*, vol. 33, no. 1, pp. 1–15, 2004.

[27] T. Cooke-Davies, "The 'real' success factors on projects," *International Journal of Project Management*, vol. 20, no. 3, pp. 185–190, 2002.

[28] J. K. Pinto and J. E. Prescott, "Variations in Critical Success Factors Over the Stages in the Project Life Cycle," *Journal of Management*, vol. 14, no. 1, pp. 5–18, 1988.

[29] D. I. Cleland, "Measuring Success: The owners viewpoint," in *Proceedings of the 18th Annual Seminar/Symposium*, pp. 6–12, Montreal, Canada, 1986.

[30] L. A. Ika, "Project success as a topic in project management journals," *Project Management Journal*, vol. 40, no. 4, pp. 6–19, 2009.

[31] S. A. Assaf and S. Al-Hejji, "Causes of delay in large construction projects," *International Journal of Project Management*, vol. 24, no. 4, pp. 349–357, 2006.

[32] A. P. C. Chan, D. Scott, and A. P. L. Chan, "Factors affecting the success of a construction project," *Journal of Construction Engineering and Management*, vol. 130, no. 1, pp. 153–155, 2004.

[33] D. W. M. Chan and M. M. Kumaraswamy, "A comparative study of causes of time overruns in Hong Kong construction projects," *International Journal of Project Management*, vol. 15, no. 1, pp. 55–63, 1997.

[34] N. R. Mansfield, O. O. Ugwu, and T. Doran, "Causes of delay and cost overruns in Nigerian construction projects," *International Journal of Project Management*, vol. 12, no. 4, pp. 254–260, 1994.

[35] M. Sambasivan and Y. W. Soon, "Causes and effects of delays in Malaysian construction industry," *International Journal of Project Management*, vol. 25, no. 5, pp. 517–526, 2007.

[36] W. Al-Ahmad, K. Al-Fagih, K. Khanfar, K. Alsamara, S. Abuleil, and H. Abu-Salem, "A taxonomy of an IT project failure: root causes," *International Management Review*, vol. 5, p. 93, 2005.

[37] J. Goedeke, M. Mueller, and O. Pankratz, "Uncovering the Causes of Information System Project Failure," 2017.

[38] L. J. May, "Major causes of software project failures," *CrossTalk*, vol. 11, pp. 9–12, 1998.

[39] P. Patanakul, "Managing large-scale IS/IT projects in the public sector: Problems and causes leading to poor performance," *The Journal of High Technology Management Research*, vol. 25, no. 1, pp. 21–35, 2014.

[40] W. Belassi and O. I. Tukel, "A new framework for determining critical success/failure factors in projects," *International Journal of Project Management*, vol. 14, no. 3, pp. 141–151, 1996.

[41] C. S. Lim and M. Z. Mohamed, "Criteria of project success: An exploratory re-examination," *International Journal of Project Management*, vol. 17, no. 4, pp. 243–248, 1999.

[42] T. Richardson, "Project management pitfalls," *Business Communications Review*, vol. 25, p. 49, 1995.

[43] I. King, "The road to continuous improvement: BPR and Project Management," *IIE Solutions*, vol. 28, pp. 22–28, 1996.

[44] G. P. Prabhakar, "What is project success: a literature review," *International Journal of Business and Management*, vol. 3, no. 9, pp. 71–79, 2009.

[45] S. M. Qureshi and C. Kang, "Analysing the organizational factors of project complexity using structural equation modelling," *International Journal of Project Management*, vol. 33, no. 1, pp. 165–176, 2015.

[46] S. Shafiei-Monfared and K. Jenab, "A novel approach for complexity measure analysis in design projects," *Journal of Engineering Design*, vol. 23, no. 3, pp. 185–194, 2012.

[47] A. P. C. Chan, D. C. K. Ho, and C. M. Tam, "Design and build project success factors: multivariate analysis," *Journal of Construction Engineering and Management*, vol. 127, no. 2, pp. 93–100, 2001.

[48] A. J. Shenhar, A. Tishler, D. Dvir, S. Lipovetsky, and T. Lechler, "Refining the search for project success factors: A multivariate, typological approach," *R&D Management*, vol. 32, no. 2, pp. 111–126, 2002.

[49] A. Tishler, D. Dvir, A. Shenhar, and S. Lipovetsky, "Identifying critical success factors in defense development projects: A multivariate analysis," *Technological Forecasting & Social Change*, vol. 51, no. 2, pp. 151–171, 1996.

[50] F. Marle and L.-A. Vidal, "Project risk management processes: Improving coordination using a clustering approach," *Research in Engineering Design*, vol. 22, no. 3, pp. 189–206, 2011.

[51] F. Marle, L.-A. Vidal, and J.-C. Bocquet, "Interactions-based risk clustering methodologies and algorithms for complex project management," *International Journal of Production Economics*, vol. 142, no. 2, pp. 225–234, 2013.

[52] L. Yang, C. Huang, and K. Wu, "The association among project manager's leadership style, teamwork and project success," *International Journal of Project Management*, vol. 29, no. 3, pp. 258–267, 2011.

[53] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological Cybernetics*, vol. 43, no. 1, pp. 59–69, 1982.

[54] T. Kohonen, "Learning vector quantization," in *Self-Organizing Maps*, pp. 175–189, Springer, 1995.

[55] J. V. Balsera, V. R. Montequin, F. O. Fernandez, and C. A. González-Fanjul, *Data Mining Applied to the Improvement of Project Management. Data Mining Applications in Engineering and Medicine*, InTech, 2012.

[56] S. G. MacDonell, "Visualization and analysis of software engineering data using self-organizing maps," in *Proceedings of the 2005 International Symposium on Empirical Software Engineering, ISESE 2005*, pp. 115–124, aus, November 2005.

[57] A. Naiem, M. El-Beltagy, and S. Seif, "Project portfolio exploration and visualization using self-organizing maps," in *Proceedings of the 10th International Conference on Informatics and Systems, INFOS 2016*, pp. 222–227, egy, May 2016.

[58] UNS, *International Standard Industrial Classification of all Economic Activities*, United Nations Statistical Office, New York, NY, USA, 1990.

[59] PMA - Determine Project Size n.d, March 2018, https://pma.doit.wisc.edu/size_factors.html.

[60] T. Kohonen, "The Self-Organizing Map," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464–1480, 1990.

[61] J. Vesanto, J. Himberg, E. Alhoniemi, and J. Parhankangas, "Self-organizing map in Matlab: the SOM Toolbox," in *Proceedings of the Matlab DSP conference*, vol. 99, pp. 16-17, 1999.

[62] J. Vesanto and E. Alhoniemi, "Clustering of the self-organizing map," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 11, no. 3, pp. 586–600, 2000.

[63] J. C. Bezdek and N. R. Pal, "Some new indexes of cluster validity," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 28, no. 3, pp. 301–315, 1998.

[64] G. W. Milligan and M. C. Cooper, "An examination of procedures for determining the number of clusters in a data set," *Psychometrika*, vol. 50, no. 2, pp. 159–179, 1985.

[65] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 2, pp. 224–227, 1978.

[66] J. Vesanto, "SOM-based data visualization methods," *Intelligent Data Analysis*, vol. 3, no. 2, pp. 111–126, 1999.

[67] K. X. Pearson, "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," *Philosophical Magazine*, vol. 50, no. 302, pp. 157–175, 2009.

Advances in Operations Research
Advances in Decision Sciences
Journal of Applied Mathematics
The Scientific World Journal
Journal of Probability and Statistics
International Journal of Mathematics and Mathematical Sciences
Hindawi
Submit your manuscripts at
www.hindawi.com
Journal of Optimization
International Journal of Engineering Mathematics
International Journal of Analysis
Journal of Complex Analysis
Advances in Numerical Analysis
Mathematical Problems in Engineering
International Journal of Differential Equations
Discrete Dynamics in Nature and Society
International Journal of Stochastic Analysis
Journal of Mathematics
Journal of Function Spaces
Abstract and Applied Analysis
Advances in Mathematical Physics